Research Note

Prepausal accuracy in the speech of adults with Down syndrome

PETER FLIPSEN Jr

University of Wisconsin–Madison, Madison, WI, USA

(Received 5 September 1997; accepted 10 March 1998)

Abstract

This was an attempt to extend one of the findings of a case report by Heselwood, Bray and Crookston, who studied the speech of an adult male with Down syndrome. Their speaker produced conversational speech more accurately and/or more intelligibly in the rhythm group immediately before a pause, suggesting that speech was being planned in units of tone groups. In this study, data from a group of six adults with Down syndrome (two males, four females) were examined to test the generalizability of that finding. As a group the six speakers produced prepausal rhythm groups significantly more intelligibly, but not with any greater degree of segmental accuracy. Possible reasons for the differences between the two studies are discussed.

Keywords: Down syndrome, speech production, speech planning, tone group, rhythm group.

Introduction

A report by Heselwood, Bray and Crookston (1995) presented speech findings on a case study of an adult male with Down syndrome (DS). Among other findings, these authors reported that their speaker produced rhythm groups (a stressed syllable and all the unstressed syllables that follow it; Cruttenden, 1986) more accurately and/or more intelligibly in prepausal position. If generalizable, this finding of improved accuracy in prepausal position has important implications for our understanding of speech planning as well as the relationships between prosodic units and segmental articulation. Drawing on Levelt’s (1989) model of speech planning, Heselwood and colleagues take the position that prosodic units (in particular the tone group) provide a link between utterance planning and speech planning. They note that such an approach provides additional insight into possible prosodic influences on speech errors (cf. Bernhardt and Stoel-Gammon, 1994; Howell and Dean, 1991; Schwartz and Goffman, 1995). Given such broad theoretical implications, and the fact that their findings were based on data from a single speaker, additional
analysis appeared warranted. The current study was an attempt to determine whether their findings on prepausal rhythm groups would be obtained with a larger group of speakers. Specifically, quantitative analysis was completed for a sample of six adult speakers with Down syndrome. The hypothesis under test was that, in adults with Down syndrome, rhythm groups are articulated more accurately and/or produced more intelligibly when they occur in prepausal position.

**Method**

Speech transcripts derived from a previously published study of the speech of 40 adults with mental retardation (Shriberg and Widder, 1990) were examined retrospectively. Included in the Shriberg and Widder study were eight speakers with DS. The aetiological status of one of the eight participants was unclear from available records, and so was not included in the current analysis. A second speaker was eliminated from the speaker pool because his speech sample included too few tone groups that met the necessary criteria for analysis (see below). Speaker characteristics are listed in table 1. Information as to the genotype of DS (i.e. mosaic vs. trisomy 21) was not available for any of the speakers. To ensure continuity between the current data and the original transcript files, published speaker numbers were retained from the original study. The resulting group included two males and four females ranging in age from 21 to 39 years. Severity of cognitive involvement was based on a review of records. The males and two of the four females were rated as moderately mentally retarded, and the two other females were rated as mildly retarded. Also shown in table 1 are the speech severity values which were based on the percentage of consonants correct (described below) from the original speech transcripts.

The original speech samples from Shriberg and Widder were taken from recordings of interviews conducted as part of a larger study of behavioural adjustment in non-institutionalized adults with mental retardation (Reynolds and Baker, 1988). The original transcriptions used the system of Shriberg and Kent (1982) and conventions for the software program PEPPER (Shriberg, 1986). In the present study the original narrow phonetic transcriptions were used for phonetic analysis with parsing of utterances adjusted to accord with conventions used by Heselwood et al. (1995).

The conversational turn served as the initial utterance unit with a turn ending when a speaker finished answering a question. Using this definition a turn did not

<table>
<thead>
<tr>
<th>Speaker no. a</th>
<th>Sex</th>
<th>Age (years)</th>
<th>Retardation level b</th>
<th>Speech severity c</th>
<th>Usable tone groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>Male</td>
<td>21</td>
<td>Moderate</td>
<td>70.7</td>
<td>13</td>
</tr>
<tr>
<td>37</td>
<td>Male</td>
<td>24</td>
<td>Moderate</td>
<td>62.0</td>
<td>18</td>
</tr>
<tr>
<td>40</td>
<td>Female</td>
<td>21</td>
<td>Mild</td>
<td>81.1</td>
<td>28</td>
</tr>
<tr>
<td>23</td>
<td>Female</td>
<td>23</td>
<td>Moderate</td>
<td>64.3</td>
<td>22</td>
</tr>
<tr>
<td>34</td>
<td>Female</td>
<td>34</td>
<td>Mild</td>
<td>82.3</td>
<td>18</td>
</tr>
<tr>
<td>9</td>
<td>Female</td>
<td>39</td>
<td>Moderate</td>
<td>73.2</td>
<td>17</td>
</tr>
</tbody>
</table>

aFrom Shriberg and Widder (1990).

bCognitive status; from available records.

cPercentage of consonants correct from original transcript.
end if the interviewer made back-channel remarks (e.g. ‘uh-hu’, ‘yes’), which occasion­ally resulted in turns in which overtalk (simultaneous talk by the speaker and examiner) occurred. Such overtalk utterances were omitted, to avoid potential transcription errors due to masking of productions or problems with tone group segmentation. Transcription conventions from PEPPER (Shriberg, 1986) include rules for disregards (words not to be included in the phonetic analysis) such as revisions, questionable transcriptions, use of filler words, and false starts. Because such words raise potential problems with both the fluency of the tone group and calculation of production accuracy, all turns containing disregards were excluded from the analysis. Remaining turns were digitized individually from the original analogue audiocassette tapes at 22 kHz, low-pass filtered at 9.8 kHz and stored using the Record utility of the software program CSpeech (Version 4; Milenkovic, 1996). Digitized waveforms were converted to spectrograms using analysing bandwidths of 300 and 400 Hz for the male and female speakers respectively.

The turns were then divided into tone groups. Although tone group boundaries may occur at moments of either open or closed juncture (Cruttenden, 1986), the current study was a test of the hypothesis that optimal accuracy occurred just prior to points of open juncture (i.e. pauses). Therefore, only tone groups immediately surrounded by pauses (open juncture tone groups) were used. Closed juncture tone groups occurred only a total of six times (no more than two per speaker); each occurrence consisting of a multi-clausal utterance without internal pauses. For current purposes a pause was operationally defined as any period of silence of at least 250 ms in duration (Miller, Grosjean and Lomanto, 1984). Due to difficulties reported with perceptual identification of pauses (cf. Stucken­enberg and O’Connell, 1988), pause durations were measured directly from the spectrograms.

Using these criteria, frequent tone group exclusions were necessary for only two speakers. A total of 17 tone groups were excluded for speaker 37 because of the frequent use of fillers such as ‘um’ and ‘uh’. For speaker 23, only the first two of numerous repetitions of the phrase ‘I dunno’ were used for the current analysis. In both cases exclusion of tone groups was motivated by the need to avoid bias in the phonetic analysis due to excess weighting on particular segmental targets.

Each of the remaining tone groups was then divided into rhythm groups (again a stressed syllable and all the unstressed syllables that follow it; Cruttenden, 1986). An independent judge with extensive experience with phonetic transcription and linguistic analysis made the stress assignments on the tone groups. The judge listened to each tone group using the Play utility of CSpeech, and was allowed to listen to each as many times as needed to assign stress status (stressed vs. unstressed) to each syllable. All tone groups for each speaker were presented sequentially as they appeared in the original transcript with the order of speaker presentation alternated quasi-randomly by gender.

Following assignment of stress status to the syllables the prepausal rhythm groups were identified. The non-prepausal portion of each was then examined to determine whether it contained any rhythm groups or consisted solely of anacrustic syllables (those occurring at the left edge of the tone group, not specifically part of a rhythm group). Tone groups in which the non-prepausal portion consisted only of anacrustic syllables were excluded from the analysis as they were not felt to represent a true test of whether rhythm group production is influenced by relative position (a prepausal rhythm group would have been matched against no rhythm groups). The result of the elimination of such tone groups was that the number of usable tone groups
ranged from five to 28 across the seven speakers. For one speaker the sample of five tone groups was judged to be too small from which to draw reasonable conclusions, and that speaker was excluded from the analysis. The number of usable tone groups for each of the six speakers in the current report is shown in the right-most column of Table 1.

Two files were created for each speaker from copies of the transcripts containing the usable tone groups. The first file consisted of the prepausal rhythm groups only, with the second containing the non-prepausal portions of the tone groups. Anacrustic syllables were included in the non-prepausal portion again, in accord with conventions used by Heselwood et al. (1995). This ensured that the entire tone group was included in the analysis, a crucial factor given the claim by Heselwood et al. of the primacy of the tone group in speech planning. The files were analysed using an upgraded version of PEPPER running in a VAX (VMS) environment. Specifically, a subroutine known as Speech Profiles (Shriberg, 1993) was used to compare prepausal rhythm groups with the non-prepausal portions for both production accuracy and intelligibility.

Reliability measures

Reliability of segmental transcription was reported in Shriberg and Widder (1990). In the current study three additional reliability estimates were obtained. The first of these involved the reliability of the II (intelligibility index; see below under Speech Metrics). A second transcriber with a similar level of experience glossed each of the six samples. II values were calculated and compared to those originally obtained. Across the six samples, differences ranged from +6.1% to −6.8% (M = +0.8%, SD = 4.7%) with an equal number of positive and negative differences. The Spearman rank-order correlation between the two sets of scores was 0.89. Given that over 10 years had passed since the original transcriptions, intra-judge reliability for II was deemed to be of limited value and therefore was not attempted.

The second reliability measure was motivated by the fact that the original transcriptions had been modified significantly, both by changing the utterance parsing to reflect turns and by eliminating tone groups. It was deemed important to assess whether the shortened versions of the transcripts continued to be representative of the speakers’ conversational speech performance. Segmental accuracy and intelligibility scores for the shortened and original versions were calculated on the speech severity metrics described below (Shriberg, Austin, Lewis, McSweeny, and Wilson, 1997a). For reliability purposes, because several of these metrics are highly intercorrelated, a subset of five was chosen. These included the three developmental consonant classes of the percentage of consonants correct (PCC), the percentage of vowels correct (PVC) and intelligibility index (II). PCC and PVC were chosen because they make the most conservative assumptions about distortions, and II was chosen because intelligibility subsumes more than segmental accuracy (Kent, Miolo and Bloedel, 1994). Across the six speakers and the five metrics, the mean difference between the original and shortened transcript was 0.0% (SD = 5.0%), with 14 of the differences positive (values for the original transcript were larger), 15 negative and one showing no difference. These data suggested that the shortened versions of the transcripts were representative of the original transcripts.

The final additional measure involved reliability of stress assignment. This was estimated by having the judge reassign stress to a randomly selected sub-sample of
35 (22%) of the tone groups, 1 week after the original assignments. A total of 84% (159/190) of the syllables received the same assignment on each occasion. Corrected for chance (Marascuilo and Levin, 1983), agreement was calculated to be 66%. Because the location of the prepausal rhythm group was central to the current analysis, reliability of identification of the final stressed syllable in the tone group was also estimated, yielding 89% exact agreement. Again corrected for chance, agreement on the last stressed syllable was 81%. To assess inter-judge reliability on stress assignment, a second judge with a similar background listened to the same subsample of 35 tone groups and assigned stress to the syllables. All assignments were compared to the original assignments of the first judge and 80% of the syllables received the same assignment (62% when corrected for chance). Agreement on the last stressed syllable was 74% (58% when corrected for chance).

Comparison metrics

The eight speech severity metrics used for the analyses were as follows (cf. Shriberg et al., 1997a):

1. The percentage of consonants correct (PCC) is the percentage of consonants produced correctly with all speech—sound deletions, substitutions and clinical distortions (cf. Shriberg, 1993, Appendix) counted as incorrect.
2. The percentage of consonants correct—adjusted (PCCA) is the percentage of consonants produced correctly, with five common clinical distortion errors counted as correct. The five distortion types are (1) labialized /l/ or /l/, (2) velarized /l/ or /l/, (3) lateralized voiced or voiceless sibilant fricatives or affricates, (4) dentalized voiced or voiceless sibilant fricatives or affricates, and (5) derhotacized /r/, /s/, or /s/. All other distortions are counted as incorrect.
3. The percentage of consonants correct—revised (PCCR) is the percentage of consonants produced correctly with all distortion errors considered correct.
4. The percentage of vowels correct (PVC) is the percentage of vowels and diphthongs produced correctly with all vowel or diphthong deletions, substitutions and distortions counted as incorrect.
5. The percentage of vowels correct—revised (PVCR) is calculated in a manner similar to the PCCR: vowel and diphthong deletions and substitutions are counted as incorrect but all distortions are counted as correct.
6. The percentage of phonemes correct (PPC) is a single metric for all segments (consonants, vowels and diphthongs), with deletions, substitutions and distortions counted as incorrect.
7. The percentage of phonemes correct—revised (PPCR) is a single metric for all segments with all distortion errors counted as correct.
8. The intelligibility index (II) is the percentage of words in the transcript for which the transcriber had provided a gloss.

For the three consonant metrics (PCC, PCCA, PCCR), the Speech Profiles program groups the 24 consonants of English into three developmental sound classes with separate percentages provided for each class: the Early-8 consonants /m,b,j,n,w,d,p,h/, the Middle-8 consonants /t,g,k,q,f,v,ʃ/ and the Late-8 consonants /l,θ,s,z, l,r,ʃ/ (Shriberg, 1993; Shriberg et al., 1997a). Early-8 percentages for the PCC and PCCA are identical since none of the common clinical distortions
scored as correct by the PCCA occur in the Early-8 group. Reporting Early-8 values for both would therefore be redundant, and will be presented here only for the PCC and PCCR.

The use of highly intercorrelated metrics such as the nested series PCC, PCCA, and PCCR, while raising Type 1 sampling concerns, was deemed justified in the current context because of the exploratory nature of the analysis. The flexibility of treating distortion errors as either correct or incorrect, as well as the ability to treat certain clinically relevant distortion types separately, appeared especially relevant for the population being studied. The cognitive delays seen in DS might motivate an expectation of a predominance of substitution and omission errors as seen in speech delay (Shriberg, Gruber and Kwiatkowski, 1994). Conversely, the adult age of these speakers might motivate an expectation of a larger proportion of distortion errors, as seen in the category residual errors (Shriberg, Austin, Lewis, McSweeny and Wilson, 1997b).

Results

Data for the six speakers were pooled and prepausal rhythm groups were compared to the non-prepausal portions of the tone groups on the 16 metrics. Results are shown in Table 2. Note that, while pooled means and standard deviations are shown

<table>
<thead>
<tr>
<th>Severity metric</th>
<th>Prepausal rhythm groups</th>
<th>Non-prepausal portion</th>
<th>Differencea</th>
<th>Tb</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCC Early-8</td>
<td>87.7 8.3</td>
<td>91.4 6.9</td>
<td>−3.7</td>
<td>2.0ns</td>
</tr>
<tr>
<td>PCCR Early-8</td>
<td>89.9 6.4</td>
<td>91.4 6.9</td>
<td>−1.5</td>
<td>5.0ns</td>
</tr>
<tr>
<td>PCC Middle-8</td>
<td>75.4 16.9</td>
<td>73.9 8.7</td>
<td>1.5</td>
<td>12.0ns</td>
</tr>
<tr>
<td>PCCA Middle-8</td>
<td>75.4 16.9</td>
<td>73.9 8.7</td>
<td>1.5</td>
<td>12.0ns</td>
</tr>
<tr>
<td>PCCR Middle-8</td>
<td>77.5 17.6</td>
<td>74.8 9.8</td>
<td>2.7</td>
<td>12.0ns</td>
</tr>
<tr>
<td>PCC Late-8</td>
<td>43.6 12.8</td>
<td>39.2 17.7</td>
<td>4.4</td>
<td>16.0ns</td>
</tr>
<tr>
<td>PCCA Late-8</td>
<td>63.5 19.0</td>
<td>59.1 17.8</td>
<td>4.4</td>
<td>14.0ns</td>
</tr>
<tr>
<td>PCCR Late-8</td>
<td>67.6 18.5</td>
<td>60.4 17.2</td>
<td>7.2</td>
<td>15.0ns</td>
</tr>
<tr>
<td>PCC Total</td>
<td>74.3 6.0</td>
<td>72.0 8.6</td>
<td>2.3</td>
<td>14.0ns</td>
</tr>
<tr>
<td>PCCA Total</td>
<td>79.9 7.5</td>
<td>77.4 7.5</td>
<td>2.5</td>
<td>14.0ns</td>
</tr>
<tr>
<td>PCCR Total</td>
<td>82.4 8.3</td>
<td>78.0 7.7</td>
<td>4.4</td>
<td>11.0ns</td>
</tr>
<tr>
<td>PVC</td>
<td>91.9 4.2</td>
<td>92.3 5.7</td>
<td>−0.4</td>
<td>8.0ns</td>
</tr>
<tr>
<td>PVCR</td>
<td>94.3 5.7</td>
<td>94.4 4.0</td>
<td>−0.1</td>
<td>10.0ns</td>
</tr>
<tr>
<td>PPC</td>
<td>80.7 4.8</td>
<td>80.9 5.3</td>
<td>−0.2</td>
<td>11.0ns</td>
</tr>
<tr>
<td>PPCR</td>
<td>86.7 7.0</td>
<td>85.2 4.7</td>
<td>1.5</td>
<td>14.0ns</td>
</tr>
<tr>
<td>II</td>
<td>97.9 4.0</td>
<td>93.5 7.4</td>
<td>4.4</td>
<td>20.0*</td>
</tr>
</tbody>
</table>

aPrepausal mean – non-prepausal mean.
bTest statistic, matched-pair Wilcoxon test (Marascuilo and Serlin, 1988).
nsNot significant (one tail; p > 0.05).
*p < 0.05.

abbreviations: PCC = Percentage of consonants correct; PCCA = Percentage of consonants correct—adjusted; PCCR = Percentage of consonants correct—revised; PVC = percentage of vowels/diphthongs correct; PVCR = percentage of vowels/diphthongs correct—revised; PPC = percentage of phonemes correct; PPCR = percentage of phonemes correct—revised; II = intelligibility index (see text for details).
for descriptive purposes, the small sample size required non-parametric inferential statistics. The matched-pair Wilcoxon test (Marascuilo and Serlin, 1988) was selected and test statistic ($T$) values are shown in the right-most column of table 2. Given the directional nature of the current hypothesis, a one-tailed test was used with an alpha level of 0.05. Although none of the purely segmental metrics yielded significant differences, the prepausal rhythm groups were significantly more intelligible than the non-prepausal portions of the tone groups. Examination of the data for the individual speakers revealed a pattern of greater intelligibility for the prepausal rhythm groups for five of the six speakers (the exception being speaker 40, in whom the difference was only 0.8% in the opposite direction).

Data across all the metrics, for individual speakers, were then examined. Because the difference scores in the analysis for individual speakers would represent the difference between single scores (and not group means), inferential statistics for each speaker on each metric were not possible. However, it was possible to examine the pattern of positive and negative differences across metrics for each speaker. As noted previously, many of the metrics used are highly intercorrelated. Therefore, a subset was selected that included the three developmental consonant classes as well as the vowels. Intelligibility index (II) was included because of the group-level findings and, as noted earlier, intelligibility is a measure reflecting more than segmental accuracy (Kent et al., 1994). Heselwood et al. also described prepausal rhythm groups as being either articulated more accurately or produced more intelligibly, suggesting the need to include metrics indexing both possibilities.

Selection of an appropriate version of the segmental metrics was made by examining the trends in group-level difference scores across the metrics. As shown in table 2, in every case there is a larger gap (or at least a less negative one) between prepausal and non-prepausal for the revised metric (i.e. where distortion errors are scored as correct). This suggested that these would provide the best opportunity for determining a difference within individual speakers, if one existed. It was for this reason that the PCCR versions of the Early-8, Middle-8 and Late-8 consonants as well as the PVCR were chosen.

Data on speaker differences between prepausal rhythm groups and the non-prepausal portions of the tone groups for the five selected metrics are provided in table 3. A one-tailed matched-pair Wilcoxon test with an alpha level of 0.05 was applied across the set of five differences for each of the six speakers, and the obtained

Table 3. Individual speaker difference data on selected speech metrics

<table>
<thead>
<tr>
<th>Speaker no.</th>
<th>PCCR Early-8</th>
<th>PCCR Middle-8</th>
<th>PCCR Late-8</th>
<th>PVCR</th>
<th>II</th>
<th>$T^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>0</td>
<td>22.5</td>
<td>20.0</td>
<td>1.5</td>
<td>1.7</td>
<td>14.5ns</td>
</tr>
<tr>
<td>37</td>
<td>-0.8</td>
<td>-16.7</td>
<td>-0.7</td>
<td>-2.4</td>
<td>7.1</td>
<td>4.0ns</td>
</tr>
<tr>
<td>40</td>
<td>-6.2</td>
<td>-2.8</td>
<td>11.2</td>
<td>2.7</td>
<td>-0.8</td>
<td>7.0ns</td>
</tr>
<tr>
<td>23</td>
<td>-5.4</td>
<td>-7.5</td>
<td>20.5</td>
<td>-2.2</td>
<td>1.6</td>
<td>6.0ns</td>
</tr>
<tr>
<td>34</td>
<td>-3.2</td>
<td>7.3</td>
<td>-2.9</td>
<td>5.1</td>
<td>4.7</td>
<td>12.0ns</td>
</tr>
<tr>
<td>9</td>
<td>6.7</td>
<td>13.2</td>
<td>-5.0</td>
<td>-5.8</td>
<td>1.6</td>
<td>10.0ns</td>
</tr>
</tbody>
</table>

$^a$Entries represent prepausal rhythm groups–non-prepausal portion differences (i.e. negative values indicate higher scores for the non-prepausal portion).

$^b$Test statistic; matched-pair Wilcoxon test.

$^c$Not significant (one-tail; $p>0.05$).
test statistic ($T$) values are shown in the right-most column. As indicated, there were no significant differences for any of the six speakers.

An analysis of rhythmic complexity was then undertaken to determine whether it had influenced the current findings. Heselwood et al. noted that, for their speaker, longer rhythm groups (units of more than two syllables in length which they term extended dynamic units) were more likely to suffer from truncation (loss of syllables) when used in non-prepausal position. It is possible that if a large proportion of the rhythm groups in the non-prepausal portions of the current samples were basic dynamic units (i.e. consisting of two or fewer syllables), there would be limited opportunity for such truncation to occur. Results indicated that 36% ($62/172$) of the rhythm groups in non-prepausal position were extended dynamic units. This was interpreted to mean that there had been sufficient opportunity for truncation to occur in non-prepausal position. In addition, a significantly smaller proportion of the rhythm groups in prepausal position (7.8%, $9/116$) were extended dynamic units ($\chi^2 = 29.84$, d.f. = 1, $p < 0.05$; Minitab Inc., 1996). This last difference suggests that if truncation were the primary mechanism for generating differences in intelligibility or segmental accuracy between prepausal and non-prepausal position, the current sample was in fact biased in favour of finding it.

It is unclear, of course, whether the impact of truncation would be to reduce intelligibility or segmental accuracy or both. One could argue that the loss of syllables would potentially make glossing of targets more difficult, thus reducing intelligibility (a possibility supported here). At the same time the loss of presumably weak or unstressed syllables would predict reduced segmental accuracy (not seen here) characterized largely by a greater proportion of omission errors. Error data were evaluated to assess this prediction and data on total number of errors by type are shown in Table 4. A chi-square test of proportions was conducted comparing omissions against substitutions plus distortions in non-prepausal position. Results indicated no significant difference ($\chi^2 = 2.42$, d.f. = 1, $p > 0.05$). Even comparing the relative proportion of omissions (against the two other error types) across non-prepausal and prepausal position, no significant difference was observed ($\chi^2 = 2.72$, d.f. = 1, $p > 0.05$). Thus, despite more opportunities for truncation to occur in non-prepausal position (i.e. because of the presence of proportionally more extended dynamic units), omission errors were not more frequent.

Heselwood et al. do not, however, contend that truncation is the only mechanism involved. They speculate that, although the segmental errors for their speaker were based in difficulties with speech prosody, the underlying problem may have been one of ‘weak initiator power’ (p. 135) based in the respiratory system. If this were the case one might postulate a higher proportion of distortion errors arising from reduced airflow in the non-prepausal portions of the tone groups. However, as with

\begin{table}  
\caption{Pooled error data\textsuperscript{a}}
\begin{tabular}{lll}
\hline
Error type & Prepausal rhythm groups & Non-prepausal portion \\
\hline
Omissions & 27 (30.3\%) & 81 (40.5\%) \\
Substitutions & 34 (38.2\%) & 70 (35.0\%) \\
Distortions & 28 (31.5\%) & 49 (24.5\%) \\
Total errors & 89 & 200 \\
\hline
\end{tabular}
\textsuperscript{a}Total number of vowel+consonant errors.
\end{table}
omissions, the distortion data for non-prepausal position shown in table 4 do not support this \( \chi^2 = 3.53 \), d.f. = 1, \( p > 0.05 \). Comparing non-prepausal to prepausal position, table 4 shows the percentages trending in the opposite direction (i.e. a larger proportion of distortions in prepausal position).

**Discussion**

The current findings provide partial support for the generalizability of the findings of Heselwood *et al.* (1995). As a group, prepausal rhythm groups in the speech of the six adults with DS were produced more intelligibly than the non-prepausal portions of the tone groups. Segmental accuracy was not significantly better.

The failure to find significant differences for segmental accuracy may, in part, be a function of differences in transcription conventions used in the two studies. The system used here, that of Shriberg and Kent (1982, 1996), relies on the assumption that the speaker’s target was an English phoneme (an assumption felt to be reasonable with speakers who had all lived at least 21 years in an English-speaking community). If such a phoneme was not realized, it is to be transcribed as a variant of the closest English target (i.e. it is assumed to be a distortion). This approach is believed to be more conservative in that the speaker is less likely to be penalized for a (phonemic) substitution and more likely to be credited with a (phonetic) distortion. Heselwood *et al.*, who used the International Phonetic Alphabet (IPA), included several instances of non-English phoneme symbols in the examples they cite, indicating that they were not making this same assumption, however. For example, they transcribed several instances of their speaker’s attempts to produce a voiced velar stop target with the IPA symbol for a voiced velar fricative. As this is not an English phoneme, such a production would be transcribed as either a frictionalized velar stop or a backed palatal fricative using the Shriberg and Kent system. Given the well-documented problems with perceptual transcription in general (Kent, 1996) and the lowered reliability of narrow phonetic transcription in particular (McSweeny and Shriberg, 1995; Shriberg and Lof, 1991; Shriberg *et al.*, 1997a), the use of different systems might account for some of the segmental differences observed between the two studies. Differences in transcription systems would, of course, be of much less importance in measures of intelligibility.

Exclusion of closed juncture tone groups from the current analysis may also have contributed to the failure to observe differences for the segmental metrics. Heselwood *et al.* cite several interesting examples from the closed juncture context. For example, their speaker frequently added a tag on to many utterances (as a separate tone group but not isolated by a pause). The presence of the tag coincided with a reduction in the accuracy of the preceding tone group segments. They speculated that, since the tag was frequently occurring and probably a highly automatized form, it required less time to process. It might then emerge before the rest of the utterance (less automatized) was fully programmed, resulting in interference. As with truncation, such interference might be realized as either a reduction in the ability to determine the gloss for the production or a distorted perception of a known target. Unfortunately, closed juncture tone groups were exceedingly rare in the current samples and thus the question of the impact of closed juncture on either segmental accuracy or intelligibility could not be examined. Differences in sampling format might have been responsible for the low frequency of closed juncture tone groups. Heselwood *et al.* (1995) used descriptions of a video in a conversation
format while the current samples were taken from interviews. The question-and-answer nature of the interviews might have resulted in shorter, or perhaps less complex, output.

The significant finding for intelligibility, despite no significant difference for segmental accuracy, may also in part reflect the fact that measurements of segmental accuracy are based on the glossable or intelligible portions of the speaker’s output and ignore the unintelligible portions, whereas intelligibility reflects the relationship between the intelligible and the unintelligible (cf. Weston and Shriberg, 1992). Thus measures of segmental accuracy and intelligibility are not derived from the same base.

It may, of course, be that the speaker described by Heselwood et al. is unique in the way in which he planned speech. According to Heselwood (personal communication, June 1997), rhythmic disturbance may be a common feature in the speech of persons with DS, but its manifestation may vary quite widely across individual speakers (see also Bray, Heselwood and Crookston, 1995; Heselwood, 1997). Individual speaker data in table 3 do suggest considerable variability within this population.

Acknowledgements

The author thanks Carol Widder for her careful transcription of the original conversational speech samples, Frederic Gruber for his patient assistance with stress assignments, and Lawrence Shriberg for his support and very insightful comments through all phases of the project. Also thanks to Robert Hagiwara, Sheryl Hall and Jane McSweeny for their assistance with reliability determination. Finally, many thanks to Barry Heselwood for his careful review of an earlier version of this manuscript. Preparation of this paper was supported by a grant from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health [R01 DC00496].

References


Prepausal accuracy in DS


MINITAB INC., 1996, Minitab Release 11 [computer program] (State College, PA: Minitab Inc.).


